Wearables in the Workplace: Sensing Interactions at the Office

Nathan Eagle, Alex (Sandy) Pentland
MIT Media Lab
nathan@media.mit.edu, sandy@media.mit.edu

Abstract

Wearable computers, such as PDAs and cell phones, are proliferating in organizations. This paper introduces a methodology for leveraging this wearable infrastructure to provide insight to the underlying social dynamics of an organization. The work lays a foundation that will enable a variety of applications, from knowledge management to organizational simulation.

1. Motivation

A wearable computing infrastructure is already inherent in many organizations. Handheld computers, otherwise known as Personal Digital Assistants (PDAs), and mobile phones have become a standard part of corporate attire. Over 10 million PDAs were sold last year, and their functionality now includes the potential for gigabytes of storage, wireless connectivity, and a processor with speeds comparable to those found in desktop PCs just a few years ago. However this wearable infrastructure is being severely underutilized. Despite the new capabilities of handheld computers, most organizations have not transcended the PalmPilot paradigm, constraining the devices to calendar and datebook functionality.

It is time for organizations to begin harnessing their wearable computing infrastructure. Hundreds of millions of dollars are spent annually on knowledge management and social network analysis - both intimately connected with face-to-face interactions, but currently quantified through surveys and mining existing corporate data repositories. This paper introduces a system that leverages the existing wearable infrastructure to quantify face-to-face interactions, to enable social network and knowledge management applications, and hopefully, to help organizations become more open and creative.

2. The Reality Mining System

The Reality Mining system is comprised of over fifty Sharp Zaurus PDAs, wireless CF cards, 256 MB SD cards, and either wired or wireless headset microphones. The wireless microphones were primarily the Jabara FreeSpeak Bluetooth headset with a 2.5mm adaptor connected to the PDA audio i/o port. Applications were written to enable the Zaurus to stream high quality audio (22 kHz, 16-bit) to an available 802.11b network, or to store the audio locally when no network is detected. Additionally, each PDA records and sends access point ID and signal strength information to a central server for location inference. Participants using wired microphones keep the PDAs in a pocket, while those with Bluetooth headsets have the option of storing the devices in a briefcase or purse.

Conversation detection and analysis algorithms were developed to process the audio streamed from the wearable. Speech detection used a variation of a multi-band center clipper over the 300 Hz to 3kHz frequency spectrum. If speech was detected from multiple people streaming audio to the same access point, waveform correlations were taken to test for a speaker's voice captured in the microphones of the others. Although a high correlation is indicative of proximity between two people, it is not a certain sign of a single conversation because the speech could be from two separate phone conversations. As shown in Figure 1 and described in [2], a single conversation contains voicing segments of one participant that are the noisy compliment of the other participants' voicing segments.

Once a conversation is detected by the wearable, a variety of conversation features can be extracted from the audio, such as speaking rate, number of interruptions, time holding the floor, volume, proximity of others, and many other turn-taking characteristics. It has been shown in the conversational analysis literature that these features are indicative of relationship [1].

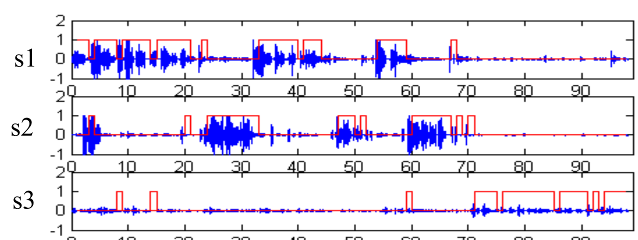

Fig. 1. A 90-second audio segment with turn-taking characteristics indicative of a single conversation.

3. Applications

Once detected, the conversation audio streams are extracted and analyzed. Table 1 shows a selection of

Proceedings of the Seventh IEEE International Symposium on Wearable Computers (ISWC'03)

features that can be gleaned from this audio data. Profiles of a participant's typical social behavior are built over time using conversation features such as speaking rate, energy, duration, participants, interruptions, transition probabilities, and time spent holding the floor. By comparing relative volume levels of a speaker's voice in multiple microphones, it even becomes possible to infer proximity of the participants to an approximate degree.

During meetings, the wearables can serve a dual purpose. Using Jon Gips' OpinionMetrics software [4], the handheld computers also enabled a user to input his or her interest level (shown in Figure 2) through an interface designed to minimize distraction.

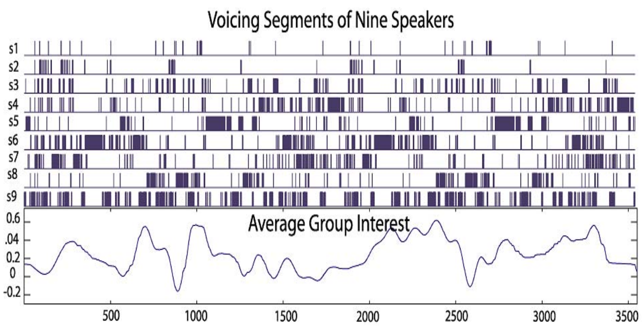

Fig. 2. A one-hour meeting with voicing segments mapped above aggregate interest level.

Speaker Number	Floor time (%)	Avg Comment (sec)	Nearest Neighbor	Transition (Name, %)	Avg Inter -est	Group Interest
s1	1.5	4.1	s8	s8-27	.21	.44
s2	2.2	2.2	s9	s9-47	.13	.36
s3	9.9	3.5	s9	s4-22	.20	.22
s4	11.4	9.6	s7	s6 -23	.05	.30
s5	12.8	8.8	s7	s9-37	.18	.33
s6	16.9	6.6	s4	s7-28	.09	.21
s7	10.1	6.6	s4	s9-30	.19	.24
s8	10.8	10.9	s1	s9-26	.40	.32
s9	24.4	6.9	s7	s6-22	.17	.25

Table 1. An analysis of the meeting from Figure 2.

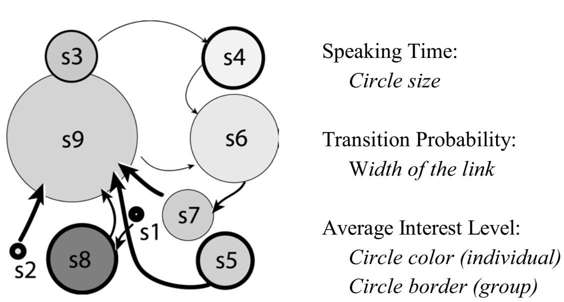

Fig. 3. A visual depiction of the group's dynamics

This system uncovers information concerning the efficiency of the group, as well as the dyadic relationships between individuals. Information from the wearable device such as the people that a user typically sits by, avoids, talks to, interrupts, and transitions to, all are correlated with relationship. In the example shown in Figure 3, a professor (s9) dominates the meeting while his advisees (s2, s7, s8) all concede the floor to him with relatively high probability - indicative of his influence [1].

4. Future Work

This work has created a foundation that will enable a variety of applications for organizations. Managers now have an opportunity to form teams around people who not only have synergistic skills sets, but also conducive social profiles. Architects have expressed interest in using this system to quantify how subtle changes to the interior of a workplace affect face-to-face communications. The measures we are currently developing to spot topics and identify questions enable a novel knowledge management system. Instead of surveys or mining static documents, this system will harness the information within face-to-face conversations by the water cooler, in a hallway, or at a meeting. Characterizing each person in an organization by local links to others, the questions they ask, and those that are posed to them, creates what is essentially a directed graph. This property allows the network to be queried for authoritative sources of information in a similar way internet search engines do today.

The Reality Mining system, comprised of a wearable platform coupled with specialized analytic software, has the potential to generate a data set never before possible in the realm of the social sciences. In anticipation of these extremely rich datasets, we are currently developing sophisticated probabilistic models that we hope will shed insight into the underlying parameters that govern group behavior [3].

6. References

[1] Atkinson, J., Heritage, J., (editors), *Structures of Social Action: Studies in Conversation Analysis*, Cambridge University Press, Cambridge, UK, 2003.

[2] Basu, S: Conversation Scene Analysis, in Dept. of EECS. Cambridge, MIT, 2002.

[3] Choudhury, T., Clarkson, B., Basu, S., and Pentland, A.: "Learning Communities: Connectivity and Dynamics of Interacting Agents" to appear in the Proceedings of IJCNN - Special Session on Autonomous Mental Development. Portland, Oregon. 2003.

[4] Devaul, R., Gips, J., Sung, M., "MIThril 2003: Applications and Architecture", to appear in the Proceedings of the 7th International Symposium on Wearable Computers. White Plains, NY. 2003.